

Identifying conserved polychaete molecular markers of metal exposure: Comparative analyses using the *Alitta virens* (Annelida, Lophotrochozoa) transcriptome

Amaia Green Etxabe, Jennifer M. Pini, Stephen Short, Luis Cunha, Peter Kille, Gordon J. Watson



PII: S1532-0456(20)30213-1

DOI: <https://doi.org/10.1016/j.cbpc.2020.108913>

Reference: CBC 108913

To appear in: *Comparative Biochemistry and Physiology, Part C*

Received date: 12 August 2020

Revised date: 7 October 2020

Accepted date: 12 October 2020

Please cite this article as: A.G. Etxabe, J.M. Pini, S. Short, et al., Identifying conserved polychaete molecular markers of metal exposure: Comparative analyses using the *Alitta virens* (Annelida, Lophotrochozoa) transcriptome, *Comparative Biochemistry and Physiology, Part C* (2018), <https://doi.org/10.1016/j.cbpc.2020.108913>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Identifying conserved polychaete molecular markers of metal exposure: comparative analyses using the *Alitta virens* (Annelida, Lophotrochozoa) transcriptome.

Amaia Green Etxabe¹, Jennifer M. Pini², Stephen Short^{3*}, Luis Cunha^{4,5}, Peter Kille³, Gordon J. Watson⁶

¹Centre for Ecology and Hydrology, Maclean Building, Benson Lane, Wallingford, Oxfordshire, UK, OX10 8BB

²CP-Texinfine (France) 60 Rue Duguesclin, 69006 Lyon, France.

³Cardiff School of Biosciences, Sir Martin Evans Building, Museum Avenue, Cardiff, UK, CF10 3AT

⁴University of Coimbra, Centre for Functional Ecology, Department of Life Sciences, Calçada Martim de Freitas, 3000-456 Coimbra, Portugal

⁵School of Applied Sciences, University of South Wales, Pontypridd, Wales, UK, CF37 4BD

⁶Institute of Marine Sciences, School of Biological Sciences, University of Portsmouth, Ferry Road, Portsmouth, UK, PO4 9LY

*Corresponding Author: E-mail: shorts@cardiff.ac.uk.

Abstract

Polychaetes are vital for evaluating the effects of toxic metals in marine systems, and sensitive molecular biomarkers should be integral to monitoring efforts. However, the few polychaete markers that exist are inconsistent, even within the same species, failing to identify gene expression changes in metal-exposed animals incurring clear metabolic costs. Comparing previously characterised polychaete metal-responsive genes with those of another carefully selected species could identify biomarkers applicable across polychaetes. The ragworm *Alitta virens* (Sars, 1835) is particularly suited for such comparisons due to its dominance of fully saline coastal areas, widespread distribution, large biomass, and its phylogenetic position relative to other polychaete 'model' resources. A transcriptome atlas for *A. virens* was generated and an RNA-seq and qPCR screening approach was used to characterise the response to chronic exposures of environmentally relevant concentrations of copper and zinc in controlled mesocosms. Genes presenting dramatic expression changes in *A. virens* were compared with known metal-responsive genes in other polychaetes to identify new possible biomarkers and assess those currently used. This revealed some current markers should probably be abandoned (e.g. *Atox1*), while others, such as *GST-Omega*, should be used with caution, as different polychaete species appear to upregulate distinct *GST-Omega* orthologues. In addition, the comparisons give some indication of genes that are induced by metal exposure across phylogenetically divergent polychaetes, including a suite of haemoglobin subunits and linker chains that could play conserved roles in metal-stress response. Although such newly identified markers need further characterisation, they offer alternatives to current markers that are plainly insufficient.

Keywords: ragworm, benthic, heavy metal, haemoglobin, polychaete, pollution

Introduction

Copper and zinc are highly toxic metals (King et al., 2004; Reish and Gerlinger, 1997) that are of great concern for ecotoxicological risk management (Luoma and Rainbow, 2008; Walker et al., 2005) and have been ranked first and third most toxic aquatic pollutants respectively (Johnson et al., 2017). Traditionally, Cu and Zn are seen as contamination legacies (Walker et al., 2005) and environmental regulators assumed levels would decline due to lower industrial inputs and legislation (Rainbow et al., 2011), however, recent analysis has shown that coastal sediment concentrations are stable or increasing (Watson et al., 2018). This is possibly due to inputs from the use of these metals in new applications, such as nanoparticles (Gao et al., 2013), or renewed use, such as replacing TBT in antifouling paints (Srinivasan and Swain, 2007). Crucially, these metals are predicted to be more bioavailable under predicted ocean acidification scenarios (Millero et al., 2009).

Polychaetes are found in vulnerable marine benthic communities and their role as contaminant vectors, combined with their ecological importance, means they are often used to analyse ecosystem health (Dean, 2008). Sub-lethal molecular biomarkers are used extensively to study the effects of pollutants on exposed invertebrates (Schirmer et al., 2010), yet the selection of biomarkers can be complex (Hook et al., 2014). For example, genetic adaptations are known to underpin the evolution of metal tolerance (Grant et al., 1989) and different species, or even populations of the same species, can evolve independent mechanisms for coping with pollutants (Reid et al., 2016). Metal exposed polychaete species studied at the molecular level represent a wide phylogenetic range (McQuillan et al., 2014; Neave et al., 2012; Rhee et al., 2012), leading to uncertainty about the universality of existing molecular biomarkers. Polychaetes (and marine Lophotrochozoa

more generally) remain underrepresented both genomically and transcriptomically (Mehr et al., 2015). A recent study illustrates the current issues regarding the lack of metal-response markers. Analysis of *H. diversicolor* suggests while exposure at a metal contaminated site induces metabolic costs that leads to reduced weight, current gene markers failed to identify any differential gene expression (Breton and Prentiss, 2019), this is despite the application of biomarkers developed using *H. diversicolor* (McQuillan et al., 2014). However, characterising transcriptomic response to metal for a carefully selected polychaete species within the context of existing analyses (McQuillan et al., 2014; Neave et al., 2012; Rhee et al., 2007a, 2007b, 2012; Won et al., 2011) could help identify reliable and sensitive cross-species markers for monitoring diverse polychaetes.

Alitta virens (M. Sars, 1835) is an ideal species for ecosystem monitoring as it inhabits marine and estuarine environments in temperate regions of the northern hemisphere. It is also the dominant polychaete (and macrofaunal species more generally) in fully saline areas and has commercial importance (Kjelsonsen and Kostka, 2005; Nielsen et al., 1995). As sediments are sinks for metals, the benthic existence of *A. virens* means it is continuously exposed to possible contaminants via sediment and porewater (Dean, 2008). *A. virens* is also phylogenetically well-placed to allow comparisons with the other polychaete transcriptomic data sets associated with metal responses; the closely related *Hediste diversicolor* and *Perinereis nuntia* (McQuillan et al., 2014; Santos et al., 2006; Rhee et al., 2012), and the more distantly related *Ophelina* (Neave et al., 2012; Rousset et al., 2007). Furthermore, toxicity experiments (Watson et al., 2018) associated with this *A. virens* transcriptomic analysis allows biomarker candidates to be anchored to measurable effects,

as recommended to provide regulators with the tools to assess environmental risk (Hook et al., 2014).

Our first aim is to generate a comprehensive *A. virens* transcriptome to provide a rich resource for ecotoxicology. Our second is to use RNASeq-qPCR screening to characterise responses of *A. virens* to chronic (3 & 6 month) exposures of environmentally relevant concentrations of Cu, Zn and Cu-Zn using sediment-spiking. Sediment-spiking enables effects to be directly linked to test chemicals (U.S.EPA, 2005). This is particularly critical as other polychaete transcriptomic studies have compared animals from clean and contaminated sites (McQuillan et al., 2014; Neave et al., 2012; Rhee et al., 2012) where responses to a metal of interest may be obscured by acclimatisation and/or the influence of a multitude of other pollutants. Overall, we aim to identify potential biomarkers for metal-induced stress applicable across diverse polychaetes that enables comparative assessment of sensitivity and tolerance. Ultimately, this will help provide sensitive and widely applicable tools to identify and monitor vulnerable ecosystems.

Materials and Methods

Mesocosm study exposure

Mesocosms and sample processing were as previously described (Watson et al., 2018). Briefly, sediment was spiked with seawater (control), $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$, ZnCl_2 or combined solutions and left for a week at 4°C in the dark. Target bioavailable concentrations were based on UK sites (Pini, 2014), resulting in seven toxicogenomic treatments: control; low Cu; low Zn; low Cu-Zn; high Cu; high Zn; high Cu-Zn. *A. virens* (1-2 g) were added to each box (3 boxes per treatment per tank [3 tanks]) with each box connected to flowing seawater. Worms were fed fish pellets, 1-2% of starting biomass twice a week. One box per treatment

per tank was destructively sampled in January and April 2013, sediment and porewater collected and worms depurated then weighed. Approximately 0.5 cm of gut tissue (~10 segments from head) was snap frozen in liquid nitrogen, homogenised in Tri-Reagent (Ambion, Life Technologies, Carlsbad, CA, USA) and stored at -80°C.

Presumed bioavailable (termed 'bioavailable' here) sediment and porewater concentrations were obtained using a BCR 3-stage sequential extraction protocol and measured by flame AAS as previously detailed (Pini, 2014). The metal distribution across three fractions were summed to give the bioavailable concentration. Recoveries for steps 1-3 were 98-111% (Cu) and 93-103% (Zn). Tissue samples were also processed with mean recovery percentages of 92% for Cu and 99.5% for Zn.

RNA isolation, library preparation and sequencing

Total RNA was extracted from the homogenised gut tissue in Tri-Reagent (see Table S1 for exposure conditions and animal numbers), purified (Direct-Zol, Zymo Research, Irvine, CA, USA), DNase I treated (New England BioLabs, Ipswich, MA, USA), concentrated (RNA Clean & Concentrator-5, Zymo Research) and quantified (NanoDrop ND-100, Thermo Fisher Scientific, Waltham, MA, USA), before RNA integrity was checked using agarose gel electrophoresis. RNA was pooled to form two libraries, these comprised of equimolar amounts of RNA from control and exposed samples and included animals from each time point and/or exposure (Table S1). Sequencing libraries were prepared and indexed for paired-end multiplexed sequencing (TruSeq Stranded mRNA Library Prep Kit, Illumina, San Diego, CA, USA). The validated libraries (Agilent BioAnalyser 2100, Agilent Technologies, Santa Clara, CA, USA) were pooled (8pM) and sequenced (100bp paired-end run with 1 lane of an Illumina HiSeq2000 flow cell v3). Raw data BCL files were converted to FASTQ files and

quality information assessed. (library preparation and sequencing was performed by Source BioScience, Nottingham, UK).

Assembly, mapping, annotation and differential expression

Raw were quality trimmed using Trimmomatic (Bolger et al., 2014) v. 0.32] using the following settings: Illuminaclip 2:30:10, leading = 3, trailing = 3 Sliding window = 4 bp, quality = 15, minimum length = 36 bp. The trimmed reads were assembled using Trinity (Grabherr et al., 2011), v. trinityrnaseq_r20140717], using the default parameters for paired-end reads with 'read normalisation', before assembled contigs were annotated using the Trinotate pipeline (Bryant et al., 2017) and a quantitative measure of transcriptome completeness was performed with BUSCO (Kriventsev et al., 2015) (v. 3, odb10 metazoa gene set). Reads and assembled transcriptome were deposited in NCBI: Bioproject No. **PRJNA627092**. Reads associated with the control and exposed samples were mapped to the transcriptome assembly [RSEM, v. 1.2.13 (Li and Dewey, 2011)]. The original counts were normalised with weighted trimmed mean of M-values (TMM) to calculate relative expression levels. As the RNASeq has no biological replication, differentially expressed transcripts were calculated using the function exactTest implemented in edgeR [v. 3.16.5 (Robinson et al., 2010)] with a common dispersion factor of 0.1. This associated a p-value with each contig sequence. The cut-off for putatively differentially expressed genes were selected experimentally using a qPCR screen. Biological and molecular function analyses were performed using Panther (Thomas et al., 2003) (v. 13.1). Enriched Gene Ontology (GO) and keyword terms associated with putatively upregulated genes were determined using DAVID (Huang et al., 2009) (v. 6.7) using the following parameters: a medium classification stringency, similarity term overlap 3, similarity threshold of 0.50, initial group membership

3, final group membership of 3, and multiple linkage threshold of 0.50. The GO terms associated with putatively up and downregulated genes was also analysed using ReviGO (Supek et al., 2011) and GOnet (Pomaznoy et al., 2018) software. For the ReviGO analysis, scored GO term lists were input (small similarity [0.5] with a SimRel semantic similarity measure). Specifically, separate scored GO term lists were created for upregulated and downregulated genes by summing fold-change levels linked to the contigs contributing to each list. GOnet analyses were performed using the log fold change values and their associated Uniprot accessions.

Orthologue determination, sequence alignment and phylogenetic tree building

Orthologous genes were determined using a Reciprocal Best Hit BLAST method (Moreno-Hagelsieb and Latimer, 2007) and sequence alignment, trimming and tree building were performed using Geneious (v. 8.1.9).

Determination of stable reference genes for qPCR validation and primer design

To develop a reliable qPCR assay, the expression of five candidate reference genes was determined using a comparative ΔC_q method (Silver et al., 2006). Briefly, a mean ΔC_q value and standard deviation (std. dv.) (R, v.3.3.0) was calculated for triplicate reactions performed for each gene in twelve independent cDNA samples (representing a range of control and exposed animals). A mean (std. dv.) was calculated to reflect the expression stability of all genes relative to the remaining four candidates (ggplot2 v. 1.0.1). The two genes presenting lowest mean std. dv. were then selected as reference genes. Putative differentially expressed genes were selected on the basis of differential expression (according to RNASeq screening) across control and exposed samples. Contig sequences representing candidate reference genes (*Elong-like*, *GAPDH-like*, *beta-actin-like*, *alpha-*

tubulin-like and *Ubiquitin-like*) and differentially expressed genes (*Unk1-like*, *Vtg-like1*, *Actin-like*, *Calbind-like*, *GST-Omega-like*, *Unk2-like* and *GST-Mu-like*) were used in conjunction with Primer3 software (Rozen and Skaletsky, 1999) (v. 4.1.0) to produce suitable primer sets for qPCR (Table S2).

cDNA production and qPCR

Total RNA (1 µg) isolated from control and exposed tissues was reverse transcribed (GO Script Promega) using OligoDT₁₅ primers. Quantitative PCR reactions (using an Eco Illumina cycler) were performed using GoTaq qPCR Master Mix (Promega, Maddison, WI, USA) and were carried out in triplicate with minus reverse transcriptase controls. Reactions were performed as follows: 95°C for 2 minutes then 45 cycles of 95°C - 5 seconds and 60°C - 30 seconds followed by a melt curve analysis. An arbitrary threshold was selected within the linear phase of the log ΔRn (fluorescence signal) against cycle number plot and the quantification cycle (Cq value) was defined as the cycle number at which ΔRn crossed the threshold. The relative expression of genes in control and exposed samples was determined using the $\Delta\Delta Cq$ method (normalising expression to the reference genes *GAPDH-like* and *Elong-like*) performed by dedicated software (EcoStudy, v. 5.0, Illumina). Amplification of the correct target sequence was confirmed by purifying (QIAquick PCR purification, Qiagen, Hilden, Germany) and sequencing (Source BioScience) all qPCR products.

Results and Discussion

Cu and Zn levels in sediments, porewater and tissues

Following exposure, the Cu/Zn levels in the sediment (bioavailable and within the porewater) and tissues from control and exposed animals was determined for representative animals (Table 1). The sediment and associated porewater concentrations

were environmentally relevant, reflecting the Sediment Quality Guideline (SQG) (Simpson et al., 2007) and global coastal contamination of sediments (Bryan and Langston, 1992; Caplat et al., 2005; Pan and Wang, 2012). The tissue concentrations were broadly comparable to levels found in polychaetes from contaminated sites (Amiard et al., 2007; Garcês and Costa, 2009; Giangrande et al., 2017; Mouneyrac et al., 2003; Neave et al., 2012; Rainbow et al., 2009) and, furthermore, Cu/Zn levels are similar to those in exposed polychaetes analysed

		Control	Low exposure	High exposure	Low CuZn	High CuZn
Copper	Nominal concentration (SQG =65, SQGH- 270)	-	70	120	70	120
	Sediment	6 ± 0.2	60 ± 8.8	25 ± 16	63 ± 6.5	81 ± 3.5
	Porewater	0.6 ± 0.2	0.9 ± 0.1	1.5 ± 0.3	1.0 ± 0.1	1.1 ± 0.1
	Tissue	9.4 ± 0.6	21 ± 4.7	79 ± 26	45 ± 13	57 ± 13
Zinc	Nominal concentration (SQG =200, SQGH- 410)	-	200	270	200	270
	Sediment	26 ± 1.4	180 ± 21	252 ± 40	163 ± 18	165 ± 23
	Porewater	2.2 ± 0.7	1.5 ± 0.5	1.6 ± 0.5	3.6 ± 0.9	3.8 ± 1.5
	Tissue	6.1 ± 7.7	108 ± 7.2	98 ± 20	88 ± 5.9	94 ± 6.4

in previous transcriptomic screening experiments (Neave et al., 2012).

Table 1. Copper and Zinc concentrations (3 and 6 months combined) in Cu, Zn and Cu-Zn mesocosm exposures. Copper and zinc mesocosm nominal bioavailable concentrations in sediment (mg kg⁻¹ dry weight) (Pini, 2014), Bioavailable sediment (mg kg⁻¹ dry weight); porewater (µg l⁻¹) and worm tissue (µg kg⁻¹ dry weight) concentrations (mean ± SEM). SQG - Sediment Quality Guideline values (mg kg⁻¹ dry sediment) SQGH - Sediment Quality Guideline High value (Simpson et al. 2007).

The *A. virens* transcriptome atlas

A representative transcriptomic atlas could be suitable for many applications. The completeness of the transcriptome can be estimated at the locus level by determining the number of near universal gene orthologues, and the extent such orthologues are present in complete or fragmented forms. The *A. virens* transcriptome atlas, produced using both control and exposed animals, was found to contain 99.2% (98.6% in complete form) of the 954 genes within the BUSCO metazoan gene set (Kriventseva et al., 2015) (transcriptome statistics in Table S3). Further analysis using Panther (Thomas et al., 2003) reveals that, despite originating from a single tissue, the sequencing depth has produced a functionally representative transcriptome when compared to the *Drosophila melanogaster* genome (Figure S1) (Thurmond et al., 2019). The data provided here will enable like-for-like quality comparisons of different data sets, such as comparative transcriptomic analysis. This valuable catalogue of polychaete genes can be used for many applications, including the development of probes and primers for gene expression analysis, particularly useful considering the increasing use of *A. virens* in evolutionary developmental biology (Kostyuchenko et al., 2015).

RNASeq and qPCR screening of *A. virens* response to Cu-Zn exposure

We provide evidence for differentially expressed genes by comparing the control and metal exposed samples. A single repeat RNASeq method in conjunction with a suitable statistical approach is thought capable of identifying differentially expressed genes with low fold change thresholds (Al Seesi et al., 2014). We have taken a more conservative approach, limiting ourselves to identifying genes with pronounced differential expression patterns, where the cut-off for differential expressed genes is decided using qPCR. Similar

approaches have provided reliable indications of pronounced gene expression differences (Al Seesi et al., 2014; Neave et al., 2012; Short et al., 2014; Wippler et al., 2016; Robertson et al., 2017). As we used qPCR to validate the RNASeq comparison of control and exposed animals, it was important to identify reliable reference genes. Expression of candidate reference genes (Table S4) in twelve samples from across the experiment (control and exposed) was determined using a comparative ΔCq method (Silver et al., 2006). The Silver et al (2006) approach has been shown comparable to other frequently used statistical approaches for determining stable reference genes (Feng et al., 2013). The stable expression suggested by the RNASeq experiment, indicated by broadly similar TMM values in control and exposed samples (Table S4) was supported by the qPCR (Figure S2, Table S5). All five genes presented a std. dev. of less than 1 and could all be used as reference genes (Svingen et al., 2015). However, *GAPDH-like* and *Elong-like* had the greatest stability (indicated by a mean std. dev of 0.4755 and 0.5117 respectively, Table S5) and were chosen for further qPCR validations.

The expression of seven genes with a range of differential expression according to RNASeq (Table S6) were analysed by qPCR, using both *GAPDH-like* and *Elong-like* as reference genes (Figure 1). The seven genes were selected on the basis of their association with a suitable range of p-values (determined using the exactTest) and because they represent a range of known and unknown functionality (Table S6). The qPCR experiment compared controls to biologically separate samples representing exposures to two concentrations of Cu, Zn and Cu-Zn for three and six months. Although variation is observed between some samples, the general expression pattern determined by qPCR agrees with the RNASeq in the p-value range $1.48e^{-11}$ to $2.7e^{-32}$. The *GST-Mu-like* gene, with the highest p-value ($3.7e^{-09}$), did not

validate well, having lower expression than suggested by RNASeq (Figure 1). Expression of both reference genes normalised to each other confirms their relatively stable expression (Figure S3). Overall, of the 12 genes validated with qPCR, 11 correlate with the RNASeq data, 5 stable (Figure S2 and S3) and 6 differentially expressed (Figure 1). This suggests that any expression change in the RNASeq associated with a p-value $\leq 1.48e^{-11}$ represents a genuine expression difference.

The genes *Unk1* and *Vtg-like1* show a pattern of clear down-regulation across both time points and all metal exposure scenarios, with the *GST-Omega-like* gene presenting an equivalent but upregulated expression pattern. This means the genes present the strikingly similar patterns across twelve biologically independent samples, with each sample representing between 11-17 animals (Table S1). Therefore, it appears these gene expression responses are triggered by Cu, Zn or a Cu-Zn combination and appear to be maintained throughout the exposure period. For the *Unk2* and the cytoskeletal-associated *Actin-like* and *Calbind-like* genes, the consistent expression between the qPCR and RNASeq only holds if expression is averaged across qPCR experiments, as there is considerable variation across samples (Figure 1). Furthermore, the expression variation does not obviously correlate with metal concentration or exposure time (Figure 1). The response variation may reflect the inconsistent relationship between metal concentrations in sediments and tissues (Table 1). It is possible that some gene responses are strikingly non-proportional to metal concentration. For some genes (in this case associated with the cytoskeleton), the environmentally relevant concentrations of Cu/Zn (Table 1) may lead to varied responses if the varying metal levels in the organisms span a threshold that must be crossed to trigger cytoskeletal modifications. This hypothesis is supported by the strikingly similar pattern of

variant expression across the equivalent qPCR samples for both cytoskeletal-associated genes. The similar expression pattern of *Unk2* and the *Actin-like* and *Calbind-like* genes (Figure 1) suggests *Unk2* may also have cytoskeletal-functions and be induced in the same threshold dependent manner (note: the *Actin-like* gene is distinct from the *beta-actin-like* gene used in reference gene analysis).

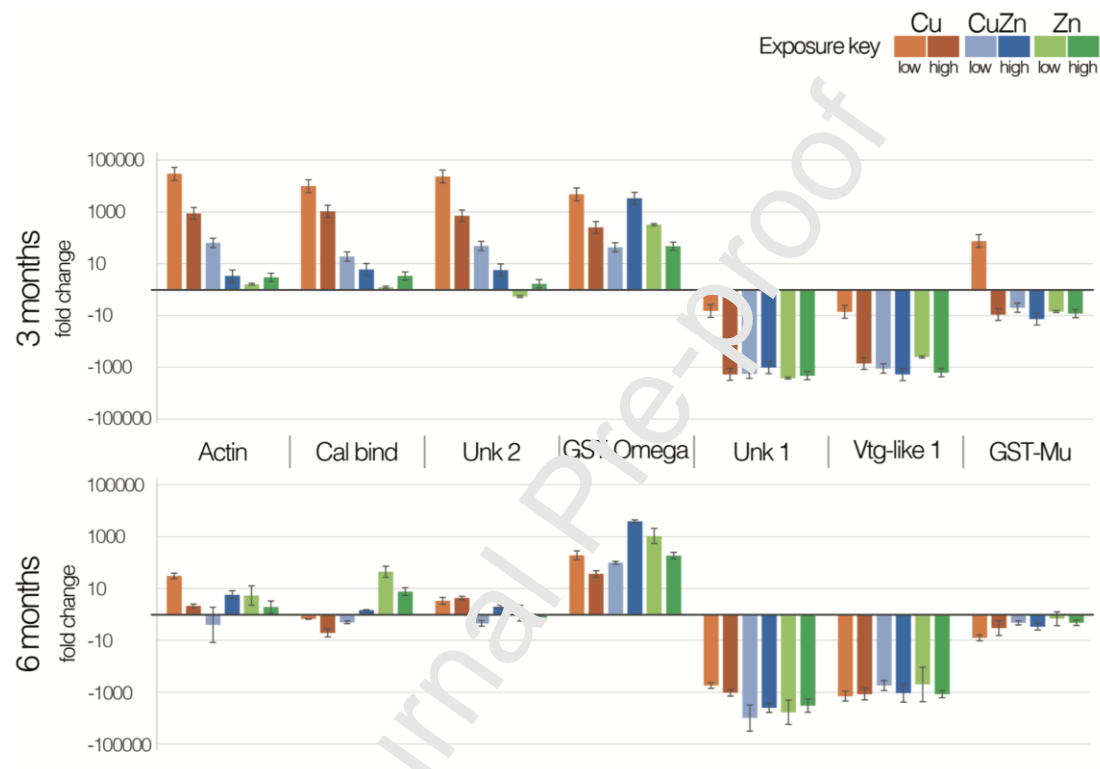


Figure 1. Comparative $\Delta\Delta C_q$ qPCR expression analysis of seven *A. virens* genes associated with a range of differential expression values suggested by RNASeq. Each column represents the average expression (3 technical repeats) of each sample relative to controls, all normalised to both reference genes (*GAPDH-like* and *Elong-like*). Samples were performed in triplicate and error bars represent the standard deviation. Expression was determined at two concentrations (low and high) of Cu, Zn and Cu-Zn for three and six months (see Table S6 for details on gene annotations and RNASeq expression).

The challenge of finding consistent biomarkers

A range of biomarkers have previously been suggested for monitoring metal exposure in polychaetes, selected for their putative roles in metal handling (McQuillan et al., 2014). However, these biomarkers have shown great variation, even within the same species (Breton and Prentiss, 2019; McQuillan et al., 2014). Critically, our findings suggest that the existing markers fail to present any notable differential expression between control and exposed animals (summarised in Table S7). For example, *Atox1* binds and transports excess intracellular Cu to secretory pathways (Kim et al., 2008). Studies of *Atox1* expression in polychaetes taken from heavily contaminated sites revealed significant but modest ~ 3x upregulation in *H. diversicolor* using qPCR (McQuillan et al., 2014), and an even greater upregulation in *Ophelina*, using RNASeq and Western blot analysis (Neave et al., 2012). However, although both *A. virens* *Atox1*-like genes possess the cysteine residues necessary for Cu binding (Rosenzweig et al., 1999), the RNASeq suggests both genes have almost identical expression in control and exposed animals. Cu handling via mechanisms requiring increased *Atox1* may have evolved independently in different species. Given these findings, *Atox1*-like genes are probably not suitable as sensitive cross-species biomarkers, despite their functional role and increased expression in some metal exposed polychaete.

Screening for biomarkers using GO term enrichment

Using the qPCR-determined cut-off, we created contig lists representing genes with apparent differential expression (DE) in exposed animals. Of the 696 and 141 contigs representing up and downregulated genes respectively, 371 and 114 could be annotated via BLAST analysis against the Uniprot database (E-value cut-off $\leq 1e-10$). The lists of Uniprot accessions were used to perform multiple analyses. Firstly, enriched Gene Ontology (GO) terms (separated into Biological Processes [BPs], Cell Components [CCs] and Molecular

Functions [MFs]) associated with the accessions were determined using DAVID (Huang et al., 2009) (Figure S4 and S5). In addition, the GO terms associated with up and downregulated genes were scored by summing fold-change levels linked to all contributing contigs. This scored GO term list was then analysed using both ReviGO (Supek et al., 2011) (Figures 2 and S6) to reveal highly expressed GO terms and GOnet software (Pomaznoy et al., 2018) (Figures S7 and S8) to reveal their various relationships.

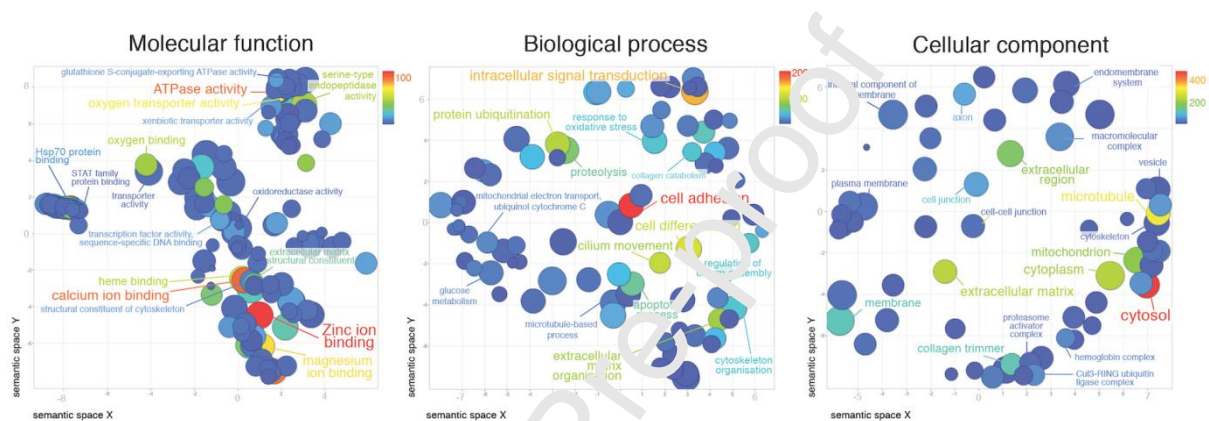


Figure 2. Summarised ReviGO analysis of 'Molecular Function', 'Biological Process' and 'Cell Component' Gene Ontology (GO) terms associated with putatively upregulated genes. GO terms associated with relevant genes were scored by summing fold-change levels linked to contributing contigs. The scored GO term list was then subsequently analysed using ReviGO.

The analysis of upregulated genes reveals clear enrichment of GO terms associated with cilia function (Figures 2, S4, S7 and S8). These are largely represented by a suite of axonemal dynein genes, with a broad range of cilia-associated genes representing the remaining GO term 'cilium'. The modest enrichment of the KOG functional group 'cell motility' in the metal-exposed polychaete *Ophelina* (Neave et al., 2012) suggests a somewhat similar response may occur across polychaetes. Assuming the modified cytoskeletal gene

expression is adaptive, it is possible to speculate on its functional role. As the intestinal epithelium of polychaetes possesses many ciliated cell types (Kermack, 1955) that come into contact with contaminated sediment, the enrichment of cilium-associated GO terms could represent attempts to replace damaged cilia. This response may also represent mechanisms that aid metal removal, as cilia in both invertebrates and vertebrates possess 'ciliary pockets' thought to be involved in vesicular trafficking (Benmerah, 2013). Indeed, gene expression suggests earthworms use a cilium-based system to sense and transport silver (Novo et al., 2015).

The analyses also provide evidence for upregulation/enrichment of GO terms broadly associated with cytoskeletal and microtubule-related functions [Figure S4 and Figure S8]). Cytoskeletal-associated genes are enriched in both *Ophelina* exposed to Cu/Zn (Neave et al., 2012) and the earthworm *Lumbricus rubellus* exposed to cadmium (Owen et al., 2008). The cytoskeleton is an early target of oxidative damage (Dalle-Donne et al., 2001) and normal reactive oxygen species (ROS) concentrations are critical for regular cytoskeletal remodelling and function (Sakai et al., 2012; Wilson and González-Billault, 2015). Therefore, the metal-exposure may lead to dysfunctional changes to the cytoskeletal composition. However, the expression alterations may be adaptive to some degree, as cytoskeletal changes can act as a signalling mechanism for oxidative stress in diverse organisms (Clark et al., 2018; Farah and Amberg, 2007) and play critical roles in the repair of damaged cells (Abreu-Blanco et al., 2012). The cytoskeletal GO terms enriched in *Ophelina* and *A. virens* are not represented by the same gene sets (Neave et al., 2012). To some extent, this likely reflects the different tissues assayed (whole body versus gut tissue). It may also reflect the evolution of divergent response mechanisms, resulting from the multi-generation field exposure of *Ophelina*

(Neave et al., 2012), compared to the exposed *A. virens* being taken from a relatively clean site. In support of this hypothesis, metal exposure of the polychaete *Hediste diversicolor*, taken from both metal-contaminated and clean sites, revealed animals originally from the contaminated-site possess more efficient strategies for sequestering and excreting metals (Mouneyrac et al., 2003). Specifically, the gut epithelium of exposed contaminated-site worms presented abundant Cu-containing lysosomes and extracellular granules, whereas exposed clean-site worms either lacked or possessed greatly reduced equivalent structures. Furthermore, while ‘spherocrystals’ were observed in the gut of both exposed clean and contaminated-site worms, these structures only act as a detoxified Zn store in the latter (Mouneyrac et al., 2003).

Enriched GO terms, associated with collagen metabolism and the extracellular matrix (Figure S4), are dominated by a series of matrix metalloproteinases (MMPs) and these genes are also components of several prominent terms in the ReviGO analysis, such as ‘Zinc ion binding’ (Figure 2). Further insight into the likely role of MMPs is given by their inclusion in the highly expressed terms ‘proteolysis’ (Figures 2) ‘proteinaceous extracellular matrix’ (Figure S4), where they are combined with other ECM-related proteins including multiple collagen α -chains. Taken together, these provide clear evidence of increased ECM remodelling in exposed animals. Collagens dominate the ECM and MMPs are critical to their degradation and modification (Lu et al., 2011), with increased MMP expression and activity observed in vertebrates exposed to oxidative stress (Alge-Priglinger et al., 2009; Siwik et al., 2001). Although knowledge of ECM remodelling is primarily based on vertebrate systems, extrapolation is reasonable as major ECM components and MMP functionalities are conserved across metazoans (Özbek et al., 2010; Page-McCaw, 2008). On this basis, the

presence of ECM and MMP GO terms may be explained if metal-induced ROS damage to intestinal cells induces repair mechanisms that involve substantial ECM remodelling. Not surprisingly, given the conservation of ECM-related genes, there is evidence of similar responses in other polychaetes. For example, even though direct comparisons of contributing genes are not possible, the KOG term 'extracellular structures' is also enriched in *Ophelina* (Neave et al., 2012) from metal contaminated sites.

Other genes with established links to responses to cell damage. For example, *E3 ubiquitin ligase* that contributes to the GO term 'protein ubiquitination' (Figure 2), has known roles in mediating response to oxidative stress (Wolyniec et al., 2013). Furthermore, the enriched 'Cul3-RING ubiquitin ligase complex' (Figure 2) is dominated by highly conserved kelch-like proteins, which function as critical components of the ubiquitin-proteasome system (Gupta and Beggs, 2014). In addition, in mammalian cells, kelch proteins bound to metals can activate a battery of genes important for cell survival against oxidative and other stresses (He and Ma, 2010).

We observe high expression of genes associated with the GO terms 'oxygen binding', 'oxygen transporter activity' and 'heme binding' (Figure 2). We also observe notable enrichment of GO terms 'oxygen transport' and 'hemoglobin complex' in the DAVID analysis (Figure S4). The term 'heme binding' includes a dual oxidase, a gene that, in vertebrates, responds to chemical-induced epithelial injury by promoting expression of genes critical to wound response, such as MMPs (van der Vliet et al., 2018). The increased expression of both dual oxidase and MMPs (see above) in exposed *A. virens* suggests a similar mechanism may be activated in the damaged gut epithelium of polychaetes. A suite of globin subunit and globin linker chains constitute the remaining genes in these GO terms. Although their

role in metal-response is uncertain (considered in more detail below), it is striking that globin subunits in *Ophelina* from metal contaminated sites also present dramatically increased expression (Neave et al., 2012).

The GO term 'ATPase activity' is associated with highly upregulated genes (Figure 2) and, although the contributing genes are functionally diverse, they all have plausible links to metal response. As well as cytoskeletal-related kinesin-like proteins, we also observe upregulation of ATP-binding cassette sub-family E member 1 (ABCE1). This highly conserved gene, critical to protein synthesis, has particularly ROS-sensitive cofactors (Fe-S clusters) (Alhebshi et al., 2012). Interestingly, increasing ABCE1 expression in yeast rescues ROS toxicity, possibly because Fe-S sequestration into ABCE1 decreases the pool of ROS-sensitive Fe-S clusters (Alhebshi et al., 2012). The upregulation of ABCE1 suggests a similar mechanism in metal-exposed *A. virens*. The Multidrug resistance-associated protein 1 (MRP1) also contributes to this GO term. MRP1 is a membrane pump that mediates the efflux of a wide variety of xenobiotics, including glutathione conjugates, and is critical to protecting cells against a range of toxic compounds (Cole, 2014). The upregulation of MRP1 in concert with specific Glutathione S-transferases (GSTs) (see discussion of *GST-Omega* below) provides a metal-detoxification/efflux pathway. Interestingly, we also see upregulation of Unconventional myosin-XIX (Myo19), an ATPase critical to the ROS-induced mitochondria localization to filopodia (Shneyer et al., 2016). Mitochondria play complex roles in stressed cells and can migrate within the cell to supply local ATP demand or provide a localized signalling platform (Shneyer et al., 2016; Tait and Green, 2012). If such mitochondrial re-localisation is occurring in *A. virens*, it may also explain some proportion of the upregulated cytoskeleton associated genes. The final contributing gene to this GO term

is Heat shock cognate 71 kDa protein (Hspa8), a highly conserved molecular chaperone implicated in the protection of the proteome from metal-induced stress (Kwon et al., 2013), that has also been found upregulated in the polychaete *Perinereis nuntia* taken from a metal contaminated site (Rhee et al., 2012).

Enrichment of specific GO terms was also observed in the relatively fewer downregulated genes, although with higher enrichment p-values (Figure S5). GO terms linked with cell junction/cell adhesion are a major feature of the DAVID and g:Profiler output (Figure S5 and S6). Downregulation of genes contributing to these groups might be related to ECM remodelling and wound-healing (supported by the upregulation MMPs and other ECM-related genes). Furthermore, ROS can disrupt cell-cell junctions between epithelial cells (Narimatsu et al., 2013) and reductions of cell junction associated proteins appears critical to normal wound-healing in *Drosophila* embryos (Hunter et al., 2015). Terms linked to ion channel/ion transport also appear in the various analyses (Figure S5 and S6). The reason for this downregulation is unclear but oxidative stress is known to modulate ion channel activity (Kiselyov and Muallem, 2016), possibly the result of considerable 'cross-talk' between REDOX and calcium signalling (Hidalgo and Donoso, 2008).

Mining for conserved and new biomarkers.

We further mined the validated gene list ($\leq 1.48 \times 10^{-11}$) to identify additional biomarkers, looking for genes known to be modulated by metal exposure in polychaetes and annelids (McQuillan et al., 2014; Mouneyrac et al., 2003; Neave et al., 2012; Novo et al., 2015; Rhee et al., 2007a, 2007b, 2012). We also include genes that have not been previously identified

as polychaete biomarkers but, given their functional roles, plausibly represent reliable and sensitive indicators of heavy metal exposure (summarised in Table 2).

Gene	Av_Gene (contig/s)	Swiss-Prot Acc.	Differential expression (DE)		Putative function	Confidence as cross-species biomarker given expression in <i>A. virens</i> and other polychaetes
			RNAseq	qPCR: low-Cu, -Zn, -Cu-Zn		
<i>GST-Omega-like</i>	c59077_g1_i1	Q6AXV9 (2e-41)	‡ ≥ 1026x ↑	~4500x, 43x, 320x ↑	GST-Omega conjugates metal ions to glutathione for excretion via ATP-dependent pumps ¹ . MRP1 is a membrane pump that mediates efflux of xenobiotics, including glutathione conjugates ² .	Moderate: <i>GST-Omega-like</i> genes up-regulated across polychaete species, but specific orthologues vary ^{3,4,18} . The co-expression of specific <i>GST-Omega</i> with MRP1 may increase confidence as a biomarker.
<i>MRP1</i>	c82198_g1_i3	P33527 (0)	‡ ≥ 56x ↑	Not tested		
<i>Vit-like</i>	c65732_g3_i1	Q90243 (3e28)	47x ↓	~7x, 1100x, 400x ↓	Defence against oxidative damage in metazoans ^{5,6} . It is thought that some Vtgs trap metal ions to protect critical cell components in <i>C. elegans</i> ⁷ .	Low: The modulation of <i>Vit-like</i> genes could be conserved across polychaetes. However, more work is needed to compare expression in other polychaete species exposed to Cu/Zn.
<i>Vit-like2</i>	c46551_g1_i1	P18948 (5e-43)	46x ↑	Not tested		
<i>Atox1-like</i>	c43190_g1_i1 c75882_g1_i1	Q9TT99 (5e-11) O08997 (1e-11)	No DE No DE	Not tested Not tested	Binds excess intracellular Cu and transports to secretory pathway ⁸ .	Very low: Large variation in response of <i>Atox1-like</i> genes to Cu/Zn between polychaete species suggests divergent regulation of these genes in response to metal.
<i>Actin-like</i>	c68928_g1_i1	Q2U7A3 (2e-81)	10643x ↑	~29,000x, 64x, 2x ↑	Unknown cytoskeletal role. Could be involved with sequestering metals into extracellular granules, observed using histology of exposed polychaetes ⁹ .	Low: Metal induced expression changes to cytoskeletal genes occur across polychaetes but specific responses between species are divergent.
<i>Calbindin-like</i>	84530_g1_i1	O18737 (2e-07)	4495x ↑	~9800x, 19x, 1.25x ↑		
Hb-subunits and *linker chains	c104830_g1_i1	P09966 (10e-82)	2970x ↑	Not tested	Unknown. Cu can change Hb-O ₂ affinity ¹⁰ or oxidise the haem group to form methemoglobin, which is unable to bind oxygen ¹¹ . We speculate that metal induced Hb/linker chains may produce erythrocytic molecules containing Hb resistant to Cu-induced conversion.	High: A specific set of haemoglobin subunits and linker chain genes is dramatically induced by Cu/Zn exposure in <i>A. virens</i> . Previous RNASeq data reveals orthologous haemoglobin subunits are induced by Cu/Zn in phylogenetically distant polychaetes ¹² . Furthermore, microarray evidence using earthworms ¹³ suggests up-regulation of specific haemoglobin subunits in response to Cu/Zn may be conserved across annelids.
	c68421_g1_i2	P02219 (2e-30)	‡ ≥ 907x ↑			
	c55755_g1_i1	P02219 (4e-43)	2074x ↑			
	c154766_g1_i1	P02220 (5e-72)	‡ ≥ 5293x ↑			
	c56032_g2_i1	P13578 (3e-83)	2879x ↑			
Dual oxidase	*c120169_g1_i1	P18207 (1e-91)	‡ ≥ 3335x ↑	Not tested	Dual oxidase responds to chemical-induced epithelial injury by promoting expression of critical to wound response/extracellular matrix remodelling, including MMPs ¹⁴ .	Moderate: Conserved genes with known roles in oxidative stress response but would need testing across phylogenetically divergent polychaetes.
	*c36018_g1_i1	P18208 (7e-128)	‡ ≥ 4463x ↑			
<i>MMP9</i>	c1858_g1_i1	O18733 (4e-28)	‡ ≥ 378x ↑	Not tested	Role in proteolysis and mediating responses to cell damage/oxidative stress ¹⁵ .	Moderate: Established association with response to cell damage but, to date, not seen in other metal exposed polychaetes.
<i>E3 Ubiquitin ligase</i>	c84381_g1_i4	Q5REG4 (3e-17)	‡ ≥ 85x ↑	Not tested	Critical to protein synthesis but has particularly Fe-S labile cofactors ¹⁶ .	Low: This highly conserved gene has ROS-labile cofactors (Fe-S clusters) ¹⁰ and increasing its expression in yeast rescues ROS toxicity, possibly via Fe-S sequestration into ABCE1 decreases the pool of sensitive Fe-S clusters ¹⁶ . Existence of the same mechanism needs establishing in a range of polychaetes.
<i>ABCE1</i>	c79906_g1_i1	P61222 (2e-96)	204x ↑	Not tested	Highly conserved molecular chaperone implicated in the protection of the proteome from metal-induced stress ¹⁷ .	Moderate: A highly conserved molecular chaperone known to help protect the proteome from metal-induced stress ¹⁷ . Possible that orthologous gene upregulated in other metal exposed polychaetes ¹⁸ but needs further testing to confirm.
<i>Hsp8</i>	c57475_g2_i2	Q90473 (2e-35)	725x ↑	Not tested	Mediates electron transfer from ubiquinol to cytochrome c, contributing to the generation of a proton gradient across the mitochondrial membrane.	Moderate: Cu/Zn exposure is associated with mis-regulation of mitochondrial-associated genes in other polychaetes ¹² and <i>MT-CYB</i> is specifically upregulated in Cu exposed earthworms ¹³ .
<i>MT-CYB</i>	c85209_g5_i1	O03553 (1e-155)	6755x ↑	Not tested		

Table 2 Summary of candidate polychaete biomarkers. Where qPCR expression is given, the values correspond to the three-month exposure. ‡ no reads were present in the control sample, the fold-change value represents a minimum estimation (estimated using pseudo-counts). Coloured shading represents genes with related functional roles, references: 1- Lu et al. 2011, 2- Cole, 2014, 3- McQuillan et al., 2014, 4- Breton and Prentiss, 2019, 5- Seehuus et al., 2006, 6- Sun and Zhang, 2015, 7- Nakamura et al., 1999, 8- Kim et al., 2008, 9- Mouneyrac et al., 2003, 10- Gutteridge, 1986, 11- Knapp et al., 2006, 12- Neave et al., 2012, 13- Bundy et al., 2008, 14- van der Vilet and Heppner, 2018, 15- Wolyniec et al., 2013, 16- Alhebshi et al., 2012, 17- Kwon et al., 2013, 18- Rhee et al., 2012).

Glutathione S-transferases (GSTs)

GSTs conjugate glutathione to xenobiotic substrates (including metals) to facilitate detoxification via ATP-dependent pumps (Ishikawa, 1992). *GST-Omega-like* genes are

associated with oxidative stress response (Kim et al., 2008) and are upregulated in several metal-exposed polychaete species, specifically *Hediste diversicolor* from metal contaminated sites (McQuillan et al., 2014), *Alitta succinea* exposed to cadmium (Won et al., 2011) and Cu-exposed *Perinereis nuntia* (Rhee et al., 2007a, 2012). Both the RNASeq and qPCR reveal upregulation of an *GST-Omega-like* gene in exposed *A. virens* (Table 2). However, sequence analysis reveals the closest *A. virens* homolog to the upregulated *A. succinea* and *P. nuntia* genes are not upregulated in *A. virens*. Furthermore, even within the same species, *GST-Omega* expression appears to vary dramatically between Cu contaminated sites (Breton and Prentiss, 2019; McQuillan et al., 2014). This divergence suggests independent evolution of *GST-Omega* response between polychaete species. Overall, the upregulation across polychaetes suggest *GST-Omega* genes play important roles in metal detoxification but, as the induced homolog varies between species, no single gene is an obvious cross-species biomarker.

Vitellogenin

Our analysis reveals opposite expression responses of different *vitellogenin-like* (*Vtg-like*) genes in exposed animals: a clear downregulation of *Vtg-like* and contrasting up-regulation of *Vtg-like2* (Table 2). It is not clear whether these expression changes represent metal-induced dysfunction or an adaptive response mechanism to mitigate the effects of exposure. However, there is enough evidence to hypothesise a functional role as, in addition to its role as a precursor of egg-yolk protein (Matozzo et al., 2008), it is clear *Vtg* also protects cells from ROS damage across metazoa (Nakamura et al., 1999; Seehuus et al., 2006; Sun and Zhang, 2015). The oxidative stress associated with metal exposure may lead to the replacement of *Vtg-like* proteins with versions that offer better protection, perhaps with increased capacities to trap metal ions (Nakamura et al., 1999). No changes to *Vtg-like*

gene expression were observed in the polychaete *Ophelina* from a Cu-Zn contaminated site (Neave et al., 2012). However, the relatively limited sequencing depth employed for *Ophelina* makes it unlikely the equivalent *Vtg* gene expression changes would be observed. Overall, further work is required to elucidate *Vtg*'s role in polychaete metal/stress response and to determine if these metal-induced expression changes play a functional and conserved role across species.

Actin-like, Calbind-like and other cytoskeletal associated genes

The large number of cytoskeletal-associated genes upregulated in metal exposed *A. virens* and *Ophelina* (Neave et al., 2012) (Figure S4) suggest these genes might be suitable biomarkers. However, the cytoskeletal-associated genes upregulated in *A. virens* and *Ophelina* are not orthologous and there is no evidence for universal markers. Furthermore, although the qPCR and RNASeq reveal upregulation of the *Actin-like* and *Calbind-like* genes in *A. virens*, the extent of upregulation is extremely variant across samples (Figure 1). The interplay of factors contributing to cytoskeletal responses is not clear and could be linked to population exposure history (Mouneyrac et al., 2003). Therefore, currently these genes are not suitable biomarkers of metal response (Table 2).

Complexes of the electron transport chain

Mitochondrial gene mis-regulation has been seen in Cu-exposed earthworms (Bundy et al., 2008) and there are established links between metal exposure and mitochondrial dysfunction (Sharpley and Hirst, 2006). The upregulation of the mitochondrial gene Cytochrome b (*CytB*) in *A. virens*, in a manner strikingly similar to that seen in Cu-exposed earthworms (Bundy et al., 2008), suggests a plausible biomarker for metal (or at least Cu) exposure (Table 2).

Haemoglobin (Hb) subunits and linker chains

The enrichment of the GO terms associated with ‘oxygen transport’ and the ‘hemoglobin complex’ (Figure S4) is predominantly due to the upregulation of multiple haemoglobin (Hb) subunits and globin linker chains. Annelids, including *A. virens*, contain giant freely dissolved ‘erythrocrucorin’ respiratory proteins (Hackert and Riggs, 2006; Weber, 1978). These proteins are assembled from multiple Hb subunits and linker chains (Knapp et al., 2006) and have been linked to polychaete metal response (e.g. Neave et al., 2012). The *Ophelina* transcriptome contained three Hb subunits that include the two paralogues *HbB1a* and *HbB1b*. Critically, while *HbB1b* was consistently expressed, *HbB1a* presented dramatic upregulation in animals from a contaminated site (Neave et al., 2012) (Figure 3 and S9). The *A. virens* transcriptome contains eleven Hb subunit genes (5 paralogue pairs) and four Hb linker-chain genes (2 paralogue pairs). These sequences likely represent distinct paralogous genes, as the amino acid identity between the most closely related sequences is ~90%, suggesting too much divergence for allelic variants. Furthermore, given the allelic homogeneity of the animals pooled for RNA samples, the sequences are unlikely to represent polymorphic alleles from divergent genotypes. Interestingly, in a manner strikingly similar to *Ophelina* (Neave et al., 2012), for each paralogue pair in *A. virens*, one is expressed consistently and the other is essentially only expressed in exposed animals (Figure 3 and Table S8). Differences in the number of Hb subunits found in *Ophelina* by in Neave et al. (2012) and *A. virens* likely results from the considerable difference in sequencing depth (~650K reads for *Ophelina* (Neave et al., 2012) and ~390 million for *A. virens*, Table S3). In Cu-exposed earthworms, upregulation of a Hb subunit in a *Lumbricus*

rubellus microarray study (Bundy et al., 2008) and increased total Hb levels in metal exposed *Lumbricus terrestris* (Calisi et al., 2011, 2013) suggest Hb upregulation may be a general annelid mechanism.

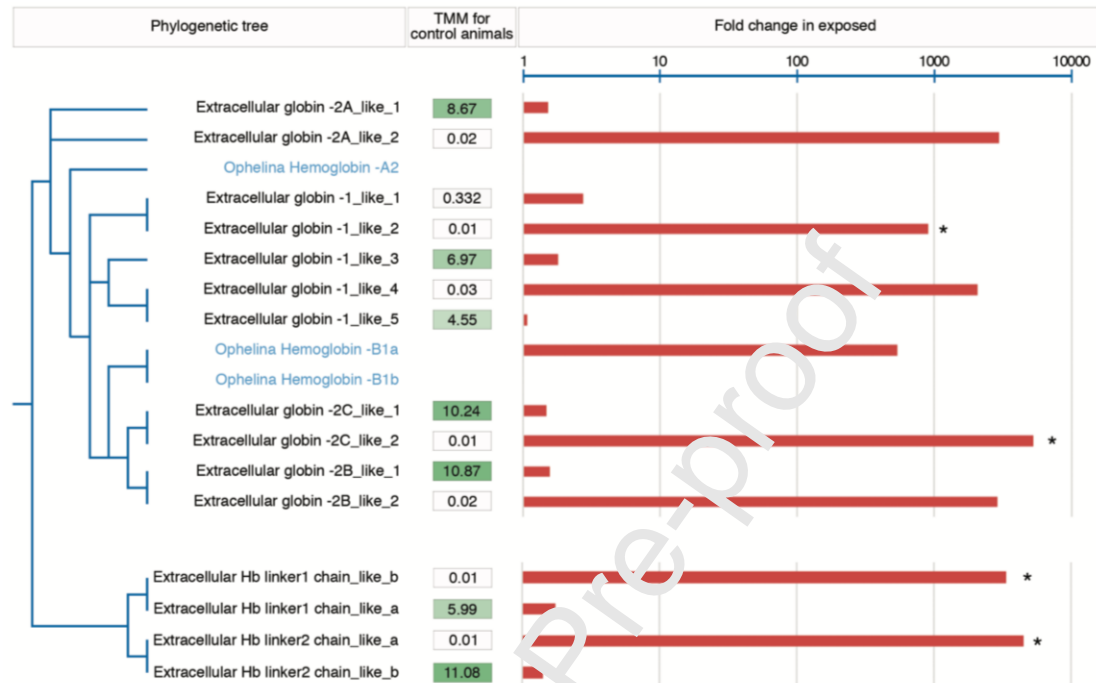


Figure 3. Haemoglobin (Hb) subunit and linker chain expression changes in the polychaetes *Alitta virens* and *Ophelina* (Neave et al. 2012) exposed to metals. Expression observed in control animals is expressed as TMM (white to dark green shows low to high expression respectively). Expression of *Ophelina* genes A2 and B1b is not published but no reported differential expression. *No reads were present in the control sample, the fold-change value represents a minimum estimation, *Ophelina* Hemoglobin -B1a is based on published read counts (Neave et al. 2012).

Assuming the upregulation of the Hb subunits and linker-chains represents an adaptive response to stress, their exact role is uncertain. However, their conserved induction across polychaetes is consistent with an important function. It is possible the upregulated subunits and linkers may act in combination with the constitutive genes to produce modified

erythrocruorins. However, the co-expression of a large repertoire of subunit and linkers could, at least in theory, lead to the production of entirely distinct erythrocruorins (Knapp et al., 2006; Weber, 1978). The subunit upregulation may reflect the need for a greater capacity to transport oxygen when under metal-induced stress, however, the negligible expression in unexposed animals (Figure 3) suggests that their presence may be counter-productive under normal conditions. Hbs isolated from the blood of polychaetes binds to toxic heavy metals (Demuyne and Dhainaut-Courtois 1993) and, despite being a critical component of annelid biology (Weber, 1978), Hbs are degraded by ROS and vulnerable to metal exposure (Gutteridge, 1986). Specifically, polychaete erythrocruorin oxygen affinity is altered by Cu/Zn (Everaarts et al., 1979) and Cu exposure can oxidise earthworm Hb to produce methemoglobin, a form that is unable to bind oxygen (Calisi et al., 2011). We propose that erythrocruorin produced using the induced genes are resistant to some combination of ROS damage, harmful changes to oxygen affinity, or Hb-methemoglobin conversion. The hypothesis that modified Hbs can mitigate toxicity is supported by evidence that free cysteines on specific Hb subunits bind hydrogen sulphide in sulphide-rich environments (Bailly et al., 2002; Chabasse et al., 2006). This capacity may extend beyond polychaetes, as a transcriptomic survey suggests that divergent Hb subunits, in the oligochaete *Olavius algarvensis*, are able to reduce the harmful effects of hydrogen sulphide produced by bacterial symbionts (Wippler et al., 2016). Further evidence comes from the polychaete *Amphitrite ornata* that possesses a dehaloperoxidase, evolved from an oxygen carrying haemoglobin, capable of breaking down volatile metabolites secreted as repellents by organisms cohabiting the same ecosystem (Lebioda et al., 1999; Barrios et al., 2014). Overall, the use of modified Hbs by polychaetes to mitigate environmental stress may be more of a rule than an exception. To determine if the Hb induction is as conserved as

suggested by the *A. virens* and *Ophelina* data, the upregulation of Hb subunits and linkers needs to be confirmed in other polychaete species. Also, Hb induction needs to be measured following exposure to specific metals and a range of other stressors, as such data should reveal insights into whether the induced subunits are critical to mitigating the specific effects of metal exposure, or play more general roles in stress response. So, although more work needs to be done, independent evidence of dramatic Hb subunit upregulation in relatively divergent polychaetes, exposed in field and mesocosm scenarios, suggests the Hb subunits could be a conserved and sensitive marker of metal exposure.

Conclusions

Comparing the Cu-Zn responses of *A. virens* with those observed in other polychaete species has confirmed that some currently used markers are inconsistent and that their use should probably be abandoned (e.g. *Atox1*). We also find that other existing markers should only be used in limited contexts, for example where particular gene orthologues only respond in specific species (e.g. *GST-Omega*). Furthermore, the comparison has identified other genes, such as haemoglobin subunits and linker chains, that appear to be metal-induced across phylogenetically divergent polychaete species. Although further expression and functional analyses of the subunits and linkers is needed to fully characterise their roles, they represent strong candidate biomarkers due to their dramatic upregulation in divergent polychaetes exposed in both mesocosm and field scenarios (Neave et al., 2012). Cu and Zn remain dominant pollutants in critical aquatic environments (Johnson et al., 2017) and polychaetes are vital for evaluating their effects (Dean, 2008). Therefore, sensitive molecular biomarkers should be central to our monitoring efforts, however, current markers fail, even in cases where metal-exposed polychaetes are incurring obvious

metabolic costs (Breton and Prentiss, 2019). Our transcriptome and comparative analysis help to fill a knowledge gap and has identified new candidate molecular markers that could help expand our ability to assess the influence of metal contamination on sensitive ecosystems.

Notes

The authors declare no competing financial interest. All sequencing data (both sequencing reads and the transcriptome assembly) will be made available through Genbank (NCBI), BioProject number: **PRJNA627092**.

Acknowledgements

This work was supported by European RDF (Interreg) grants IVA 3C and CHRONEXPO.

References

- Abreu-Blanco, M.T., Watts, J.J., Verjoo, J.M. and Parkhurst, S.M., 2012. Cytoskeleton responses in wound repair. *Cellular and Molecular Life Sciences*, 69, 2469-2483.
- Al Seesi, S., Tiagueu, Y.T., Zelikovsky, A., Măndoiu, I.I., 2014. Bootstrap-based differential gene expression analysis for RNA Seq data with and without replicates. *BMC Genomics* 15, S2.
- Alge-Priglinger, C.S., Kreutzer, T., Obholzer, K., Wolf, A., Mempel, M., Kernt, M., Kampik, A., Priglinger, S.G., 2009. Oxidative stress-mediated induction of MMP-1 and MMP-3 in human RPE cells. *Investigative Ophthalmology & Visual Science* 50, 5495-5503.
- Alhebshi, A., Sideri, T.C., Holland, S.L., Avery, S.V., 2012. The essential iron-sulfur protein Rli1 is an important target accounting for inhibition of cell growth by reactive oxygen species. *Molecular Biology of the Cell* 23, 3582-3590.

- Amiard, J.-C., Geffard, A., Amiard-Triquet, C., Crouzet, C., 2007. Relationship between the lability of sediment-bound metals (Cd, Cu, Zn) and their bioaccumulation in benthic invertebrates. *Estuarine, Coastal and Shelf Science* 72, 511-521.
- Bailly, X., Jollivet, D., Vanin, S., Deutsch, J., Zal, F., Lallier, F. and Toulmond, A., 2002. Evolution of the sulfide-binding function within the globin multigenic family of the deep-sea hydrothermal vent tubeworm *Riftia pachyptila*. *Molecular Biology and Evolution*, 19, 1421-1433.
- Barrios, D.A., D'Antonio, J., McCombs, N.L., Zhao, J., Franzen, S., Schmidt, A.C., Sombers, L.A. and Ghiladi, R.A., 2014. Peroxygenase and oxidase activities of dehaloperoxidase-hemoglobin from *Amphitrite ornata*. *Journal of the American Chemical Society* 136, 7924-7925.
- Benmerah, A., 2013. The ciliary pocket. *Current opinion in cell biology* 25, 78-84.
- Bolger, A.M., Usadel, B., Lohse, M., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114-2120.
- Breton, T.S., Prentiss, N.K., 2019. Metal stress-related gene expression patterns in two marine invertebrates, *Hediste diversicolor* (Annelida, Polychaeta) and *Littorina littorea* (Mollusca, Gastropoda), at a former mining site. *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology* 225, 108588.
- Bryan, G., Langston, W., 1992. Bioavailability, accumulation and effects of heavy metals in sediments with special reference to United Kingdom estuaries: a review. *Environmental Pollution* 76, 89-131.
- Bryant, D.M., Johnson, K., DiTommaso, T., Tickle, T., Couger, M.B., Payzin-Dogru, D., Lee, T.J., Leigh, N.D., Kuo, T.-H., Davis, F.G., 2017. A tissue-mapped axolotl *de novo* transcriptome enables identification of limb regeneration factors. *Cell Reports* 18, 762-776.
- Bundy, J.G., Sidhu, J.K., Rana, F., Spurgeon, D.J., Svendsen, C., Wren, J.F., Stürzenbaum, S.R., Morgan, A.J., Kille, P., 2008. 'Systems toxicology' approach identifies coordinated metabolic responses to copper in a terrestrial non-model invertebrate, the earthworm *Lumbricus rubellus*. *BMC Biology* 6, 25.

- Calisi, A., Lionetto, M.G., Sanchez-Hernandez, J.C., Schettino, T., 2011. Effect of heavy metal exposure on blood haemoglobin concentration and methemoglobin percentage in *Lumbricus terrestris*. *Ecotoxicology* 20, 847.
- Calisi, A., Zaccarelli, N., Lionetto, M.G. and Schettino, T., 2013. Integrated biomarker analysis in the earthworm *Lumbricus terrestris*: application to the monitoring of soil heavy metal pollution. *Chemosphere*, 90, 2637-2644.
- Caplat, C., Texier, H., Barillier, D., Lelievre, C., 2005. Heavy metals mobility in harbour contaminated sediments: the case of Port-en-Bessin. *Marine Pollution Bulletin* 50, 504-511.
- Chabasse, C., Bailly, X., Rousselot, M. and Zal, F., 2006. The multigenic family of the extracellular hemoglobin from the annelid polychaete *Arenicola marina*. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, 144, 319-325.
- Clark, M.S., Thorne, M.A.S., King, M., Hipperson, P., Hoffman, J.I., Peck, L.S., 2018. Life in the intertidal: Cellular responses, methylation and energetics. *Functional Ecology* 32, 1982-1994.
- Cole, S.P., 2014. Multidrug resistance protein 1 (MRP1, ABCC1), a “multitasking” ATP-binding cassette (ABC) transporter. *Journal of Biological Chemistry* 289, 30880-30888.
- Dalle-Donne, I., Rossi, R., Milzani, A., Di Simplicio, P., Colombo, R., 2001. The actin cytoskeleton response to oxidants: from small heat shock protein phosphorylation to changes in the redox state of actin itself. *Free Radical Biology and Medicine* 31, 1624-1632.
- Dean, H.K., 2008. The use of polychaetes (Annelida) as indicator species of marine pollution: a review. *Revista de Biología Tropical* 56, 11-38.
- Demuyne, S. and Dhainaut-Courtois, N., 1993. Identification of extracellular haemoglobin as the major high molecular weight cadmium-binding protein of the polychaete annelid *Nereis diversicolor*. *Comparative Biochemistry and Physiology Part C: Pharmacology, Toxicology and Endocrinology*, 106, 467-472.

- Everaarts, J., Osborne, W., Verhaaf, E., 1979. Effect of copper and zinc on some oxygen binding properties of the haemoglobin of the polychaete *Arenicola marina*. *Netherlands Journal of Sea Research* 13, 571-580.
- Farah, M.E., Amberg, D.C., 2007. Conserved actin cysteine residues are oxidative stress sensors that can regulate cell death in yeast. *Molecular Biology of the Cell* 18, 1359-1365.
- Feng, L., Yu, Q., Li, X., Ning, X., Wang, J., Zou, J., Zhang, L., Wang, S., Hu, J., Hu, X. and Bao, Z., 2013. Identification of reference genes for qRT-PCR analysis in Yesso Scallop *Patinopecten yessoensis*. *PloS One* 8, p.e75609.
- Gao, Y., Luo, Z., He, N., Wang, M.K., 2013. Metallic nanoparticle production and consumption in China between 2000 and 2010 and associative aquatic environmental risk assessment. *Journal of Nanoparticle Research* 15, 1681.
- Garcês, J., Costa, M.H., 2009. Trace metals in populations of *Marphysa sanguinea* (Montagu, 1813) from Sado estuary: effect of body size on accumulation. *Scientia Marina* 73, 605-616.
- Giangrande, A., Licciano, M., del Pasqua, M., Faruzzi, F.P., Migoni, D., Stabili, L., 2017. Heavy metals in five Sabellidae species (Annelida, Polychaeta): ecological implications. *Environmental Science and Pollution Research* 24, 3759-3768.
- Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, C., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29, 644.
- Grant, A., Hateley, J.G., Jones, N.V., 1989. Mapping the ecological impact of heavy metals on the estuarine polychaete *Nereis diversicolor* using inherited metal tolerance. *Marine Pollution Bulletin* 20, 235-238.
- Gupta, V.A., Beggs, A.H., 2014. Kelch proteins: emerging roles in skeletal muscle development and diseases. *Skeletal Muscle* 4, 11.
- Gutteridge, J.M., 1986. Iron promoters of the Fenton reaction and lipid peroxidation can be released from haemoglobin by peroxides. *FEBS letters* 201, 291-295.

- Hackert, M.L., Riggs, A.F., 2006. When size matters. *Structure* 14, 1094-1096.
- He, X., Ma, Q., 2010. Critical cysteine residues of Kelch-like ECH-associated protein 1 in arsenic sensing and suppression of nuclear factor erythroid 2-related factor 2. *Journal of Pharmacology and Experimental Therapeutics* 332, 66-75.
- Hidalgo, C., Donoso, P., 2008. Crosstalk between calcium and redox signaling: from molecular mechanisms to health implications. *Antioxidants & Redox Signaling* 10, 1275-1312.
- Hook, S.E., Gallagher, E.P., Batley, G.E., 2014. The role of biomarkers in the assessment of aquatic ecosystem health. *Integrated Environmental Assessment and Management* 10, 327-341.
- Huang, D.W., Sherman, B.T., Zheng, X., Yang, J., Imamichi, T., Stephens, R., Lempicki, R.A., 2009. Extracting biological meaning from large gene lists with DAVID. *Current protocols in bioinformatics* 27:13.11.
- Hunter, M.V., Lee, D.M., Harris, T.J., Fernandez-Gonzalez, R., 2015. Polarized E-cadherin endocytosis directs actomyosin remodeling during embryonic wound repair. *Journal of Cell Biology* 210, 801-816.
- Ishikawa, T., 1992. The ATP-dependent glutathione S-conjugate export pump. *Trends in Biochemical Sciences* 17, 463-468.
- Johnson, A.C., Donnachie, R.L., Sutherland, J.P., Jürgens, M.D., Moeckel, C., Pereira, M.G., 2017. An alternative approach to risk rank chemicals on the threat they pose to the aquatic environment. *Science of the Total Environment* 599, 1372-1381.
- Kermack, D.M., 1955. The anatomy and physiology of the gut of the polychaete *Arenicola marina* (L.). *Proceedings of the Zoological Society of London* 125, 347-381.
- Kim, B.-E., Nevitt, T., Thiele, D.J., 2008. Mechanisms for copper acquisition, distribution and regulation. *Nature chemical biology* 4, 176.
- King, C., Dowse, M., Simpson, S., Jolley, D., 2004. An assessment of five Australian polychaetes and bivalves for use in whole-sediment toxicity tests: toxicity and accumulation of copper and zinc from water and sediment. *Archives of Environmental Contamination and Toxicology* 47, 314-323.
- Kiselyov, K., Muallem, S., 2016. ROS and intracellular ion channels. *Cell Calcium* 60, 108-114.

- Knapp, J.E., Pahl, R., Šrajcar, V., Royer, W.E., 2006. Allosteric action in real time: Time-resolved crystallographic studies of a cooperative dimeric hemoglobin. *Proceedings of the National Academy of Sciences* 103, 7649-7654.
- Kostyuchenko, R.P., Kozin, V.V., Filippova, N.A., Sorokina, E.V., 2019. FoxA expression pattern in two polychaete species, *Alitta virens* and *Platynereis dumerilii*: Examination of the conserved key regulator of the gut development from cleavage through larval life, postlarval growth, and regeneration. *Developmental Dynamics*. 248, 728–743.
- Kristensen, E., Kostka, J., 2005. Macrofaunal Burrows and Irrigation in Marine Sediment: Microbiological and Biogeochemical. *Interactions between Macro and Microorganisms in Marine Sediments*. Washington, DC: American Geophysical Union, 115-157.
- Kriventseva, E.V., Zdobnov, E.M., Simão, F.A., Ioannidis, P., Waterhouse, R.M., 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210-3212.
- Kwon, J.Y., Weon, J.-I., Koedrith, P., Park, K.-S., Kim, I.S., Seo, Y.R., 2013. Identification of molecular candidates and interaction networks via integrative toxicogenomic analysis in a human cell line following low-dose exposure to the carcinogenic metals cadmium and nickel. *Oncology reports* 30, 1185-1194.
- Lebioda, L., LaCount, M.W., Zhang, E., Chen, Y.P., Han, K., Whitton, M.M., Lincoln, D.E. and Woodin, S.A., 1999. An enzymatic globin from a marine worm. *Nature* 401, 445-445.
- Li, B., Dewey, C.N., 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC bioinformatics* 12, 323.
- Lu, P., Takai, K., Weaver, V.M., Werb, Z., 2011. Extracellular matrix degradation and remodeling in development and disease. *Cold Spring Harbor perspectives in biology* 3, a005058.
- Luoma, S.N., Rainbow, P.S., 2008. Metal contamination in aquatic environments: science and lateral management. Cambridge University Press.

- Matozzo, V., Gagné, F., Marin, M.G., Ricciardi, F., Blaise, C., 2008. Vitellogenin as a biomarker of exposure to estrogenic compounds in aquatic invertebrates: a review. *Environment international* 34, 531-545.
- McQuillan, J.S., Kille, P., Powell, K., Galloway, T.S., 2014. The regulation of copper stress response genes in the polychaete *Nereis diversicolor* during prolonged extreme copper contamination. *Environmental Science & Technology* 48, 13085-13092.
- Mehr, S., Verdes, A., DeSalle, R., Sparks, J., Pieribone, V., Gruber, D.F., 2015. Transcriptome sequencing and annotation of the polychaete *Hermodice carunculata* (Annelida, Amphinomidae). *BMC Genomics* 16, 1-13.
- Millero, F.J., Woosley, R., Ditrolio, B., Waters, J., 2009. Effect of ocean acidification on the speciation of metals in seawater. *Oceanography* 22, 72-85.
- Moreno-Hagelsieb, G., Latimer, K., 2007. Choosing BLAST options for better detection of orthologs as reciprocal best hits. *Bioinformatics* 24, 319-324.
- Mouneyrac, C., Mastain, O., Amiard, J.C., Amiara-Triquet, C., Beaunier, P., Jeantet, A.-Y., Smith, B.D., Rainbow, P.S., 2003. Trace-metal detoxification and tolerance of the estuarine worm *Hediste diversicolor* chronically exposed in their environment. *Marine Biology* 143, 731-744.
- Nakamura, A., Yasuda, K., Adachi, H., Sakurai, Y., Ishii, N., Goto, S., 1999. Vitellogenin-6 is a major carbonylated protein in aged nematode, *Caenorhabditis elegans*. *Biochemical and Biophysical Research Communications* 264, 580-583.
- Narimatsu, T., Ozawa, Y., Miyake, S., Kubota, S., Hirasawa, M., Nagai, N., Shimmura, S., Tsubota, K., 2013. Disruption of cell-cell junctions and induction of pathological cytokines in the retinal pigment epithelium of light-exposed mice. *Investigative Ophthalmology & Visual Science* 54, 4555-4562.
- Neave, M.J., Streten-Joyce, C., Nouwens, A.S., Glasby, C.J., McGuinness, K.A., Parry, D.L., Gibb, K.S., 2012. The transcriptome and proteome are altered in marine polychaetes (Annelida) exposed to elevated metal levels. *Journal of Proteomics* 75, 2721-2735.

Nielsen, A.M., Eriksen, N.T., Iversen, J.L., Riisgård, H.U., 1995. Feeding, growth and respiration in the polychaetes *Nereis diversicolor* (facultative filter-feeder) and *N. virens* (omnivorous)-a comparative study. Marine Ecology Progress Series 125, 149-158.

Novo, M., Lahive, E., Díez-Ortiz, M., Matzke, M., Morgan, A.J., Spurgeon, D.J., Svendsen, C., Kille, P., 2015. Different routes, same pathways: Molecular mechanisms under silver ion and nanoparticle exposures in the soil sentinel *Eisenia fetida*. Environmental Pollution 205, 385-393.

Owen, J., Hedley, B.A., Svendsen, C., Wren, J., Jonker, M.J., Hankard, P.K., Lister, L.J., Stürzenbaum, S.R., Morgan, A.J., Spurgeon, D.J., 2008. Transcriptome profiling of developmental and xenobiotic responses in a keystone soil animal, the oligochaete annelid *Lumbricus rubellus*. BMC Genomics 9, 266.

Özbek, S., Balasubramanian, P.G., Chiquet-Ehrismann, P., Tucker, R.P., Adams, J.C., 2010. The evolution of extracellular matrix. Molecular Biology of the Cell 21, 4300-4305.

Page-McCaw, A., 2008. Remodeling the model organism: matrix metalloproteinase functions in invertebrates, Seminars in cell & developmental biology. Elsevier, 14-23.

Pan, K., Wang, W.-X., 2012. Trace metal contamination in estuarine and coastal environments in China. Science of the Total Environment 421, 3-16.

Pini, J., 2014. An assessment of the impacts of chronic exposure of copper and zinc on the polychaete *Nereis (Alitta) virens* using an integrated ecotoxicological approach. University of Portsmouth.

Pomaznoy, M., Ha, B., Peters, B., 2018. GOnet: a tool for interactive Gene Ontology analysis. BMC Bioinformatics 19, 470.

Rainbow, P., Kriefman, S., Smith, B., Luoma, S., 2011. Have the bioavailabilities of trace metals to a suite of biomonitors changed over three decades in SW England estuaries historically affected by mining? Science of the Total Environment 409, 1589-1602.

- Rainbow, P., Smith, B., Luoma, S., 2009. Differences in trace metal bioaccumulation kinetics among populations of the polychaete *Nereis diversicolor* from metal-contaminated estuaries. *Marine Ecology Progress Series* 376, 173-184.
- Reid, N.M., Proestou, D.A., Clark, B.W., Warren, W.C., Colbourne, J.K., Shaw, J.R., Karchner, S.I., Hahn, M.E., Nacci, D., Oleksiak, M.F., Crawford, D.L., Whitehead, A., 2016. The genomic landscape of rapid repeated evolutionary adaptation to toxic pollution in wild fish. *Science* 354, 1305-1308.
- Reish, D.J., Gerlinger, T.V., 1997. A review of the toxicological studies with polychaetous annelids. *Bulletin of Marine Science* 60, 584-607.
- Rhee, J.S., Lee, Y.M., Hwang, D.S., Lee, K.W., Kim, I.C., Shin, K.H., Raisuddin, S., Lee, J.S., 2007a. Molecular cloning and characterization of omega class glutathione S-transferase (GST-O) from the polychaete *Neanthes succinea*: biochemical comparison with theta class glutathione S-transferase (GST-T). *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology* 146, 471-477.
- Rhee, J.S., Lee, Y.M., Hwang, D.S., Won, E.J., Raisuddin, S., Shin, K.H. and Lee, J.S., 2007b. Molecular cloning, expression, biochemical characteristics, and biomarker potential of theta class glutathione S-transferase (GST-T) from the polychaete *Neanthes succinea*. *Aquatic toxicology*, 83, 104-115.
- Rhee, J.S., Won, E.J., Kim, R.O., Choi, B.S., Choi, I.Y., Park, G.S., Shin, K.H., Lee, Y.M. and Lee, J.S., 2012. The polychaete, *Perinereis auntia* ESTs and its use to uncover potential biomarker genes for molecular ecotoxicological studies. *Environmental Research*, 112, 48-57.
- Robertson, L.S., Galbraith, H.S., Iwanowicz, D., Blakeslee, C.J. and Cornman, R.S., 2017. RNA sequencing analysis of transcriptional change in the freshwater mussel *Elliptio complanata* after environmentally relevant sodium chloride exposure. *Environmental Toxicology and Chemistry* 36, 2352-2366.
- Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139-140.
- Rosenzweig, A.C., Huffman, D.L., Hou, M.Y., Wernimont, A.K., Pufahl, R.A., O'Halloran, T.V., 1999. Crystal structure of the Atx1 metallochaperone protein at 1.02 Å resolution. *Structure* 7, 605-617.

- Rousset, V., Pleijel, F., Rouse, G.W., Erséus, C., Siddall, M.E., 2007. A molecular phylogeny of annelids. *Cladistics* 23, 41-63.
- Rozen, S., Skaletsky, H., 1999. Primer3 on the WWW for General Users and for Biologist Programmers, in: S. Misener, S.A. Krawetz (Eds.), *Bioinformatics Methods and Protocols*. Humana Press, Totowa, NJ, 365-386.
- Sakai, J., Li, J., Subramanian, K.K., Mondal, S., Bajrami, B., Hattori, H., Jia, Y., Dickinson, B.C., Zhong, J., Ye, K., 2012. Reactive oxygen species-induced actin glutathionylation controls actin dynamics in neutrophils. *Immunity* 37, 1037-1049.
- Santos, C.S.G., Pleijel, F., Lana, P., Rouse, G.W., 2006. Phylogenetic relationships within Nereididae (Annelida:Phyllodocida). *Invertebrate Systematics* 19, 557-575.
- Schirmer, K., Fischer, B.B., Madureira, D.J., Pillai, S., 2010. Transcriptomics in ecotoxicology. *Analytical and Bioanalytical Chemistry* 397, 917-923.
- Seehuus, S.-C., Norberg, K., Gimsa, U., Kneking, T., Amdam, G.V., 2006. Reproductive protein protects functionally sterile honey bee workers from oxidative stress. *Proceedings of the National Academy of Sciences of the United States of America* 103, 962-967.
- Sharpley, M.S., Hirst, J., 2006. The inhibition of mitochondrial complex I (NADH: ubiquinone oxidoreductase) by Zn^{2+} . *Journal of Biological Chemistry* 281, 34803-34809.
- Shneyer, B.I., Ušaj, M., Henn, A., 2016. Myo19 is an outer mitochondrial membrane motor and effector of starvation-induced filopodia. *Journal of Cell Science* 129, 543-556.
- Short, S., Yang, G., Guler, Y., Green Etxabe, A., Kille, P. and Ford, A.T., 2014. Crustacean intersexuality is feminization without demasculinization: implications for environmental toxicology. *Environmental Science & Technology* 48, 13520-13529.
- Silver, N., Best, S., Jiang, J., Thein, S.L., 2006. Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. *BMC Molecular Biology* 7, 33.
- Simpson, S., Batley, G., Chariton, A., 2007. Revision of the ANZECC/ARMCANZ sediment quality guidelines.

- Siwik, D.A., Pagano, P.J., Colucci, W.S., 2001. Oxidative stress regulates collagen synthesis and matrix metalloproteinase activity in cardiac fibroblasts. *American Journal of Physiology-Cell Physiology* 280, C53-C60.
- Srinivasan, M., Swain, G.W., 2007. Managing the use of copper-based antifouling paints. *Environmental Management* 39, 423-441.
- Sun, C., Zhang, S., 2015. Immune-relevant and antioxidant activities of vitellogenin and yolk proteins in fish. *Nutrients* 7, 8818-8829.
- Supek, F., Bošnjak, M., Škunca, N., Šmuc, T., 2011. REVIGO summarizes and visualizes long lists of gene ontology terms. *PloS One* 6.
- Svingen, T., Letting, H., Hadrup, N., Hass, U., Vinggaard, A.M., 2015. Selection of reference genes for quantitative RT-PCR (RT-qPCR) analysis of rat tissues under physiological and toxicological conditions. *PeerJ* 3, e855.
- Tait, S.W., Green, D.R., 2012. Mitochondria and cell signalling. *Journal of Cell Science* 125, 807-815.
- Thomas, P.D., Campbell, M.J., Kejariwal, A., Mi, H., Karlak, B., Daverman, R., Diemer, K., Muruganujan, A., Narechania, A., 2013. PANTHER: a library of protein families and subfamilies indexed by function. *Genome Research* 13, 2129-2141.
- Thurmond, J., Goodman, J.L., Strelets, V.B., Attrill, H., Gramates, L.S., Marygold, S.J., Matthews, B.B., Millburn, G., Antonazzo, C., Trovisco, V. and Kaufman, T.C., 2019. FlyBase 2.0: the next generation. *Nucleic Acids Research* 47, D759-D765.
- van der Vliet, A., Danyal, K., Heppner, D.E., 2018. Dual oxidase: a novel therapeutic target in allergic disease. *British Journal of Pharmacology* 175, 1401-1418.
- Walker, C.H., Sibly, R., Hopkin, S.P., Peakall, D.B., 2005. *Principles of ecotoxicology*. CRC press.
- Watson, G.J., Pini, J.M., Richir, J., 2018. Chronic exposure to copper and zinc induces DNA damage in the polychaete *Alitta virens* and the implications for future toxicity of coastal sites. *Environmental Pollution* 243, 1498-1508.
- Weber, R.E., 1978. Respiratory pigments, *Physiology of annelids*. Academic Press, 393-446.

Wilson, C., González-Billault, C., 2015. Regulation of cytoskeletal dynamics by redox signaling and oxidative stress: implications for neuronal development and trafficking. *Frontiers in Cellular Neuroscience* 9, 381.

Wippler, J., Kleiner, M., Lott, C., Gruhl, A., Abraham, P.E., Giannone, R.J., Young, J.C., Hettich, R.L., Dubilier, N., 2016. Transcriptomic and proteomic insights into innate immunity and adaptations to a symbiotic lifestyle in the gutless marine worm *Olavius algarvensis*. *BMC Genomics* 17, 942.

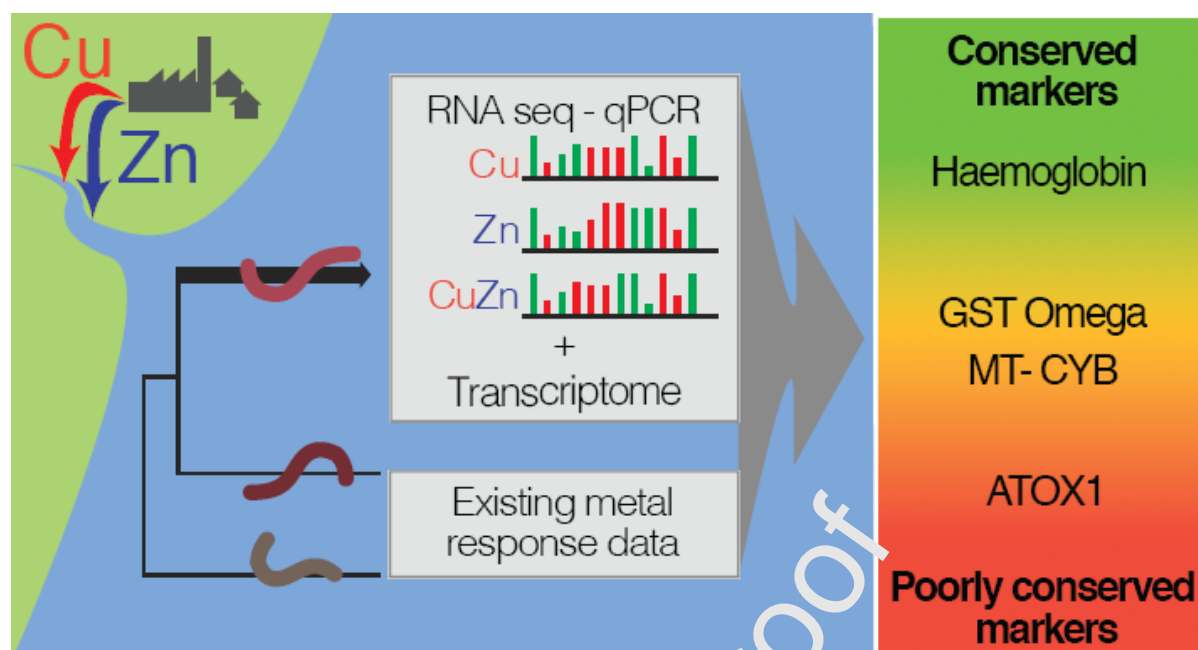
Wolyniec, K., Levav-Cohen, Y., Jiang, Y., Haupt, S., Haupt, Y., 2013. The E6AP E3 ubiquitin ligase regulates the cellular response to oxidative stress. *Oncogene* 32, 3510–3519.

Won, E.-J., Kim, R.-O., Rhee, J.-S., Park, G.S., Lee, J., Shin, K.-H., Lee, Y.-M., Lee, J.-S., 2011. Response of glutathione S-transferase (GST) genes to cadmium exposure in the marine pollution indicator worm, *Perinereis nuntia*. *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology* 154, 82-92.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Journal Pre-proof



Graphical abstract

Highlights

- We provide a comprehensive transcriptome for the polychaete *Alitta virens*.
- Metal-induced genes across divergent polychaetes are identified.
- Current markers should be abandoned (*Atox1*) or used conservatively (*GST-Omega*).
- Haemoglobins induced in divergent metal-exposed polychaetes may be better markers.